JOURNAL OF MULTIVARIATE ANALYSIS **24**, 252–264 (1988)

Weak Convergence to the Matrix Stochastic Integral $\int_0^1 B\, dB'$

P. C. B. PHILLIPS*

Cowles Foundation for Research in Economics, Yale University

Communicated by the Editors

The asymptotic theory of regression with integrated processes of the ARIMA type frequently involves weak convergence to stochastic integrals of the form $\int_0^1 W\, dW$, where $W(r)$ is standard Brownian motion. In multiple regressions and vector autoregressions with vector ARIMA processes, the theory involves weak convergence to matrix stochastic integrals of the form $\int_0^1 B\, dB'$, where $B(r)$ is vector Brownian motion with a non-scalar covariance matrix. This paper studies the weak convergence of sample covariance matrices to $\int_0^1 B\, dB'$ under quite general conditions. The theory is applied to vector autoregressions with integrated processes.

1. INTRODUCTION

Let $\{y_t\}_0^\infty$ be a multiple (n-vector) time series generated by

$$y_t = Ay_{t-1} + u_t; \qquad t = 1, 2, ... \tag{1}$$

$$A = I_n; \tag{2}$$

$$y_0 = \text{random with a certain fixed distribution.} \tag{3}$$

Under very general conditions on the sequence of innovations, $\{u_t\}_1^\infty$, in (1), $t^{-1/2}y_t$ converges almost surely to standardized vector Brownian motion $t^{-1/2}B(t)$ on $C^n[0, \infty]$. The covariance matrix, Ω, of $B(t)$ depends on the serial covariance properties of $\{u_t\}_1^\infty$. If the sequence $\{u_t\}_1^\infty$ is

Received July 18, 1986.

AMS 1980 subject classifications: primary 62E20, secondary 62M10.

Key words and phrases: integrated process, invariance principle, near integrated time series, stochastic integral, vector autoregression, weak convergence.

* This paper was written while the author was a Monash University Visiting Professor during May/June 1986. I am grateful to members of the Department of Econometrics and Operations Research for their hospitality and to Monash University for its support during this visit. My thanks also go to Glena Ames for her skill and effort in typing the manuscript of this paper and to the National Science Foundation for research support under Grant SES 8519595.

0047-259X/88 $3.00

stationary with spectral density matrix $f_{uu}(\lambda)$ satisfying $f_{uu}(0) > 0$ ("$>$" here signifies positive definite) then $\Omega = 2\pi f_{uu}(0)$. Strong invariance principles of this type have been proved recently by Berkes and Philipp [1] and by Eberlain [4].

Weak invariance principles follow directly from these strong convergence results as shown by Philipp and Stout [8]. In this case, it is usual to define the partial sum process $S_t = \sum_1^t u_j$ and construct the following random element of $D^n[0, 1]$:

$$X_T(r) = T^{-1/2} S_{[Tr]} = T^{-1/2} S_{j-1}; \qquad (j-1)/T \leqslant r < j/T.$$

Then as $T \uparrow \infty$,

$$X_T(r) \Rightarrow B(r), \tag{4}$$

where $B(r)$ is vector Brownian motion on $C^n[0, 1]$, with covariance matrix Ω. In (4) we use the symbol "$\Rightarrow$" to signify weak convergence of the associated probability measures. Billingsley [2] provides an extensive discussion of such weak invariance principles in the scalar case ($n = 1$) and gives many useful applications.

One major time series application of (4) is to the theory of regression for integrated processes. If $\{u_t\}_1^\infty$ is generated by a linear process such as a finite-order stationary and invertible vector ARMA model then y_t is known as an integrated process of order one (Box and Jenkins [3]). We are often interested in the asymptotic behavior of statistics from linear least-squares regressions with integrated processes. Thus, from the first-order vector autoregression of y_t on y_{t-1} in (1) we obtain the regression coefficient matrix

$$\hat{A} = \left(\sum_1^T y_t y_{t-1}'\right)\left(\sum_1^T y_{t-1} y_{t-1}'\right)^{-1}.$$

Here, $\hat{A}$ is a simple function of the sample moments of y_t. To the extent that y_t behaves asymptotically like vector Brownian motion, we might expect the asymptotic behavior of $\hat{A}$ to be described by a corresponding functional of Brownian motion.

To be more precise consider standardized deviations of $\hat{A}$ about I_n:

$$T(\hat{A} - I) = \left(T^{-1} \sum_1^T u_t y_{t-1}'\right)\left(T^{-2} \sum_1^T y_{t-1} y_{t-1}'\right)^{-1}. \tag{5}$$

By simple calculations we may write the sample second moment $T^{-2} \sum_1^T y_{t-1} y_{t-1}'$ as a quadratic functional of the random element $X_T(r)$, at least up to a term of $o_p(1)$. That is,

$$T^{-2} \sum_1^T y_{t-1} y_{t-1}' = \int_0^1 X_T(r)\, X_T(r)'\, dr + o_p(1).$$

Result (4) and the continuous mapping theorem then establish that

$$T^{-2}\sum_{1}^{T} y_{t-1}y'_{t-1} \Rightarrow \int_0^1 B(r)\,B(r)'\,dr \tag{6}$$

as $T\uparrow\infty$.

In a similar way, we might expect that

$$T^{-1}\sum_{1}^{T} u_t y'_{t-1} \Rightarrow \int_0^1 dB(r)\,B(r)' \tag{7}$$

at least when $\{u_t\}_1^\infty$ is a sequence of square integrable martingale differences. However, unlike (6), (7) cannot be obtained by a simple application of (4) and the continuous mapping theorem. The reason is that we cannot write the sample covariance matrix $T^{-1}\sum_1^T u_t y'_{t-1}$ as a continuous functional of the random element $X_T(r)$. Moreover, the limit process $\int_0^1 dBB'$ (we shall sometimes suppress the argument of the random function in integrals of this type) is a matrix stochastic integral and, since $B(r)$ is almost surely (vector Wiener measure) of unbounded variation, $\int_0^1 dBB'$ cannot be considered as the (mean square) limit of a Riemann Stieltjes sum. Furthermore, when the innovations u_t are not martingale differences, $E(u_t y'_{t-1})\neq 0$, in general, and there is no reason to expect (7) to hold.

In the scalar case ($n=1$, $A=a$, $\Omega=\omega^2$) the problems described in the previous paragraph are easily resolved. We simply write

$$S_T^2=\sum_1^T u_t^2+2\sum_2^T\left(\sum_1^{t-1}u_s\right)u_t$$

and, then, under quite general conditions as $T\uparrow\infty$,

$$\begin{aligned} T^{-1}\sum_1^T y_{t-1}u_t &= \frac{1}{2}\left\{(T^{-1/2}X_T(1))^2 - T^{-1}\sum_1^T u_t^2\right\}+o_p(1)\\ &\Rightarrow \frac{1}{2}\{\omega^2 W(1)^2-\omega_0^2\}, \end{aligned}\tag{8}$$

where $W(r)$ denotes standard Brownian motion on $C[0,1]$ and where $\omega_0^2=\lim_{T\to\infty}T^{-1}\sum_1^T E(u_t^2)$. Here $T^{-1}\sum_1^T u_t^2\to\omega_0^2$ a.s. by a suitable strong law for weakly dependent time series (e.g., McLeish [7]). In view of the formula $\int_0^1 W\,dW=\frac{1}{2}(W(1)^2-1)$ and the fact that $\omega W(r)\equiv B(r)$ (here the symbol "$\equiv$" signifies equality in distribution) we deduce that

$$T^{-1}\sum_1^T y_{t-1}u_t \Rightarrow \int_0^1 B\,dB+\frac{1}{2}(\omega^2-\omega_0^2). \tag{9}$$

This reduces to the formula suggested above in (7) iff $\omega^2=\omega_0^2$.

Thus, in the scalar case, we obtain the following limit law for the autoregressive coefficient:

$$T(\hat{a}-1) \Rightarrow \left\{\int_0^1 B\,dB + \frac{1}{2}(\omega^2 - \omega_0^2)\right\} \Big/ \left\{\int_0^1 B(r)^2\,dr\right\}. \tag{10}$$

Equation (10) is proved in Phillips [10] and it generalizes the simple formula $\int_0^1 B\,dB/\int_0^1 B^2\,dr$ that was first suggested by White [14] for the case where the innovation sequence is i.i.d. $N(0, \omega^2)$.

When $n > 1$ the argument that was used above to deduce (9) no longer applies. In fact, partial summation of the outer product $S_T S_T'$ yields

$$S_T S_T' = \sum_1^T u_t u_t' + \sum_2^{T-1} \left(\sum_1^{t-1} u_s\right) u_t' + \sum_2^{T-1} u_t \left(\sum_1^{t-1} u_s\right)'$$

so that, in place of (8), we now obtain

$$\begin{aligned} T^{-1} \sum_1^T (y_{t-1} u_t' + u_t y_{t-1}') &= X_T(1)\, X_T(1)' - T^{-1} \sum_1^T u_t u_t' + o_p(1) \\ &\Rightarrow B(1)\, B(1)' - \Omega_0, \end{aligned} \tag{11}$$

where $\Omega_0 = \lim_{T\to\infty} T^{-1} \sum_1^T E(u_t u_t')$. Determination of the limit law of the matrix $T^{-1} \sum_1^T y_{t-1} u_t'$ is not possible from (11), although the joint limiting distribution of its diagonal elements may be deduced. However, the latter is insufficient for many problems of central interest, such as the limiting distribution of the regression coefficients (5).

The purpose of the present paper is to obtain the matrix analogue of (9) directly. Our approach permits a wide class of possible innovation sequences and our main result is directly applicable to the study of regression statistics such as (5). It should also be useful in other contexts where weak convergence to the matrix stochastic integral $\int_0^1 B\,dB'$ is needed. Some econometric examples are given in Phillips and Durlauf [13].

2. Main Results

We shall require $\{u_t\}_1^\infty$ to satisfy conditions which are sufficient to ensure the validity of (4). In particular, we impose:

Assumption 2.1. (a) $E(u_t) = 0$ all t;

(b) $\sup_{i,t} E\,|u_{it}|^{\beta+\varepsilon} < \infty$ for some $\beta > 2$ and $\varepsilon > 0$;

(c) $\Omega = \lim_{T\to\infty} T^{-1} E(S_T S_T')$ exists and is positive definite;

(d) $\{u_t\}_1^\infty$ is strong mixing with mixing numbers α_m that satisfy

$$\sum_1^\infty \alpha_m^{1-2/\beta} < \infty. \tag{12}$$

If $\{u_i\}_1^\infty$ is weakly stationary then (c) is, in fact, implied by the mixing condition (d) (Theorem 18.5.3 of Ibragimov and Linnik [6]). In this case, we obtain

$$\begin{aligned}\Omega &= E(u_1 u_1') + \sum_{k=2}^{\infty} E(u_1 u_k') + \sum_{k=2}^{\infty} E(u_k u_1') \\ &= \Omega_0 + \Omega_1 + \Omega_1', \qquad \text{say.} \end{aligned} \tag{13}$$

Under Assumption 2.1 we have:

LEMMA 2.2. *If* $\{u_t\}_1^\infty$ *is a sequence of random n-vectors that satisfy Assumption* 2.1 *then as* $T \uparrow \infty$, $X_T(r) \Rightarrow B(r)$, *vector Brownian motion with covariance matrix* Ω.

It is convenient to introduce a multiple $(n \times 1)$ time series $\{z_t(x)\}_1^\infty$ generated by the model

$$z_t(x) = F z_{t-1}(x) + u_t; \qquad t = 1, 2, ... \tag{14}$$

$$F = \exp\{(x/T)G\} \tag{15}$$

$$z_0(x) = y_0. \tag{16}$$

Here x is a scalar and G is an arbitrary $n \times n$ matrix. When $x = 0$ the model is equivalent to (1)–(3). Note that as $T \uparrow \infty$, $F \to I_n$ so that for fixed $x \neq 0$, $z_t(x)$ behaves, at least asymptotically, like an integrated process. Such processes were introduced in Phillips [9, 12] and were called *near integrated time series*. Note also that since F depends on T, (14) in fact defines a triangular array of near integrated time series $\{\{z_{tT}(x)\}_{t=1}^{T}\}_{T=1}^{\infty}$. We will suppress the additional subscript T on z_{tT} to simplify notation.

Back substitution in (14) yields

$$z_t(x) = \sum_{j=1}^{t} \exp\{((t-j)x/T)G\}u_j + \exp\{(tx/T)G\}y_0. \tag{17}$$

Define

$$\begin{aligned}\dot{z}_t &= (d/dx)\, z_t(x) \\ &= G \sum_{1}^{t} \exp\{((t-j)\, x/T)G\}((t-j)/T)u_j + G(t/T)\exp\{(tx/T)G\}y_0. \end{aligned} \tag{18}$$

We now consider the asymptotic behavior of sample moments of these processes.

LEMMA 2.3. *If* $\{u_t\}_1^\infty$ *satisfies Assumption* 2.1 *and* $\{z_t(x)\}_1^\infty$ *is a near integrated time series generated by* (14)–(16), *then as* $T \uparrow \infty$,

(a) $T^{-1}\dot{z}_T(x)z_T(x) \Rightarrow GL_G(1, x)\,K_G(1, x)'$;

(b) $T^{-2}\sum_1^T \dot{z}_t z_t(x)' \Rightarrow G\int_0^1 L_G(r, x)\,K_G(r, x)'\,dr$;

(c) $T^{-2}\sum_1^T z_t(x)\,z_t(x)' \Rightarrow \int_0^1 K_G(r, x)\,K_G(r, x)'\,dr$,

where

$$K_G(r, x) = \int_0^r \exp((r-s)\,xG)\,dB(s)$$

$$L_G(r, x) = \int_0^r \exp\{(r-s)\,xG\}(r-s)\,dB(s).$$

We also need:

LEMMA 2.4. *If $B(r)$ is vector Brownian motion with covariance matrix Ω and $J_c(r) = \int_0^r \exp\{(r-s)C\}\,dB(s)$ then*

$$J_c(1)\,J_c(1)' = \Omega + c\int_0^1 J_c(r)\,J_c(r)'\,dr + \int_0^1 J_c(r)\,J_c(r)'\,drC'$$
$$+\int_0^1 J_c(r)\,dB(r)' + \int_0^1 dB(r)\,J_c(r)' \tag{19}$$

for any $n \times n$ matrix C.

LEMMA 2.5. *If $\{u_t\}_1^\infty$ satisfies Assumption 2.1 and $\{z_t(x)\}_1^\infty$ is a near integrated process generated by (14)–(16), then as $T \uparrow \infty$,*

(a) $T^{-1}\sum_1^T \{\dot{z}_{t-1}(x)\,u_t' + u_t\dot{z}_{t-1}(x)'\} \Rightarrow G\int_0^1 L_G(r, x)\,dB(r)' + \int_0^1 [dB(r)\,L_G(r, x)']G'$;

(b) $T^{-1}\sum_1^T \dot{z}_{t-1}(x)u_t' \Rightarrow G\int_0^1 L_G(r, x)\,dB(r)'$;

(c) $T^{-1}\sum_1^T z_{t-1}(l)\,u_t' - T^{-1}\sum_1^T y_{t-1}u_t' \Rightarrow \int_0^1 K_G(r, l)\,dB(r)' - \int_0^1 B(r)\,dB(r)'$.

We are now in a position to establish our main result:

THEOREM 2.6. *If $\{u_t\}_1^\infty$ is weakly stationary and satisfies Assumption 2.1 and if $\{y_t\}_0^\infty$ is generated by (1)–(3), then as $T \uparrow \infty$,*

(a) $T^{-1}\sum_1^T y_{t-1}u_t' \Rightarrow \int_0^1 B(r)\,dB(r)' + \Omega_1$;

(b) $T^{-1}\sum_1^T z_{t-1}(l)u_t' \Rightarrow \int_0^1 K_G(r, l)\,dB(r)' + \Omega_1$,

where

$$\Omega_1 = \lim_{T\to\infty} T^{-1}\sum_1^T E(y_{t-1}u_t') = \sum_{k=2}^\infty E(u_1u_k').$$

COROLLARY 2.7. *If $\{u_t\}_1^\infty$ is a sequence of stationary martingale differences that satisfy Assumption* 2.1 *and if $\{y_t\}_0^\infty$ is generated by* (1)–(3), *then as $T\uparrow\infty$,*

$$T^{-1}\sum_1^T y_{t-1}u_t' \Rightarrow \int_0^1 B(r)\,dB(r)'.$$

Theorem 2.6 may be extended to include sequences $\{u_t\}_1^\infty$ which are not weakly stationary with some strengthening of the moment and mixing conditions (b) and (d) of Assumption 2.1. The details are not given here since the case of predominant interest is that of weakly stationary innovations in (1).

We may now deduce the relevant asymptotics for regression statistics such as (5). In particular, we have:

THEOREM 2.8. *If the conditions of Theorem* 2.6 *hold then as $T\uparrow\infty$,*

$$T(\hat{A}-1) \Rightarrow \left\{\int_0^1 B(r)\,dB(r)' + \Omega_1\right\}'\left\{\int_0^1 B(r)\,B(r)'\,dr\right\}^{-1}. \tag{20}$$

Note that in the scalar case (setting $\Omega_1 = \omega_1$) we have $\omega^2 = \omega_0^2 + 2\omega_1$ and (20) reduces to the earlier formula (10).

3. PROOFS

Proof of Lemma 2.2. See Phillips [11].

Proof of Lemma 2.3. To prove (a) we note that

$$\begin{aligned}
T^{-1/2}z_T(x) &= \sum_{j=1}^T \exp\{((1-j/T)x)G\}\int_{(j-1)/T}^{j/T} dX_T(s) + O_p(T^{-1/2}) \\
&= \sum_{j=1}^T \int_{(j-1)/T}^{j/T} \exp\{(1-s)\,xG\}\,dX_T(s) + O_p(T^{-1/2}) \\
&= \int_0^1 \exp\{(1-s)\,xG\}\,dX_T(s) + O_p(T^{-1/2}) \\
&\Rightarrow \int_0^1 \exp\{(1-s)\,xG\}\,dB(s), \qquad \text{as} \quad T\uparrow\infty.
\end{aligned}$$

in view of Lemma 2.2 and the continuous mapping theorem. In a similar way we find that

$$T^{-1/2}\dot{z}_T(x) \Rightarrow G\int_0^1 \exp\{(1-s)\,xG\}(1-s)\,dB(s)$$

and result (a) follows directly. To prove (b) we write

$$
\begin{aligned}
T^{-2}\sum_1^T \dot{z}_t z_t'
&= T^{-2}\sum_{i=1}^{T}\left[G\sum_{j=1}^{i}\exp\{((i-j)/T)\,xG\}((i-j)/T)u_j\right] \\
&\quad\cdot\left[\sum_{k=1}^{i} u_k'\exp\{((i-k)/T)\,xG'\}\right]+O_p(T^{-1/2}) \\
&= \sum_{i=1}^{T}\int_{(j-1)/T}^{i/T} dr\left[G\sum_{j=1}^{i}\int_{(j-1)/T}^{j/T}\exp\{(r-s)\,xG\}(r-s)\,dX_T(s)\right] \\
&\quad\cdot\left[\sum_{k=1}^{i}\int_{(k-1)/T}^{k/T} dX_T(t)'\exp\{(r-t)\,xG'\}\right]+O_p(T^{-1/2}) \\
&= \int_0^1\int_0^r\int_0^r G\exp\{(r-s)\,xG\}(r-s)\,dX_T(s)\,dX_T(t')\exp\{(r-t)\,xG'\}\,dr \\
&\quad+O_p(T^{-1/2}) \\
&\Rightarrow \int_0^1\left[G\int_0^r \exp\{(r-s)\,xG\}(r-s)\,dB(s)\right] \\
&\quad\times\left[\int_0^r dB(t)'\exp\{(r-t)\,xG'\}\right]dr \\
&= G\int_0^1 L_G(r,x)\,K_G(r,x)'\,dr
\end{aligned}
$$

as required. Part (c) follows in a similar fashion.

Proof of Lemma 2.4. First define $\xi(r)=\int_0^r \exp(-sC)\,dB(s)$ and note that $J_c(r)=\exp(rC)\,\xi(r)$. Now by the multivariate Ito formula for stochastic differentiation we have

$$d\{\xi(r)\,\xi(r)'\}=d\xi(r)\,\xi(r)'+\xi(r)\,d\xi(r)'+\exp(-rC)\,\Omega\exp(-rC)'\,dr.$$

Hence

$$
\begin{aligned}
&\int_0^1 [\exp(rC)\,d\{\xi(r)\,\xi(r)'\}\exp(rC')] \\
&\qquad=\int_0^1 dB(r)\,J_c(r)'+\int_0^1 J_c(r)\,dB(r)'+\Omega,
\end{aligned}
$$

leading to the result as stated.

Proof of Lemma 2.5. From (14) we obtain

$$\begin{aligned} z_t(x)\, z_t(x)' - z_{t-1}(x)\, z_{t-1}(x)' &= (xG)\, T^{-1} z_{t-1}(x)\, z_{t-1}(x)' + T^{-1} z_{t-1}(x)\, z_{t-1}(x)'(xG') \\ &\quad + z_{t-1}(x)\, u_t' + u_t z_{t-1}(x)' + u_t u_t' + O_p(T^{-1}) \end{aligned}$$

and averaging over t we find

$$\begin{aligned} &T^{-1} z_T(x)\, z_T(x)' \\ &\quad = (xG)\, T^{-2} \sum_1^T z_{t-1}(x)\, z_{t-1}(x)' + T^{-2} \sum_1^T z_{t-1}(x)\, z_{t-1}(x)'(xG) \\ &\qquad + T^{-1} \sum_1^T z_{t-1}(x) u_t' + T^{-1} \sum_1^T u_t z_{t-1}(x)' + T^{-1} \sum_1^T u_t u_t' + O_p(T^{-1}). \end{aligned}$$

Differentiating with respect to x yields

$$\begin{aligned} &T^{-1} \dot{z}_T z_T' + T^{-1} z_T \dot{z}_T' \\ &\quad = xGT^{-2} \sum_1^T (\dot{z}_{t-1} z_{t-1}' + z_{t-1} \dot{z}_{t-1}') \\ &\qquad + T^{-2} \sum_1^T (\dot{z}_{t-1} z_{t-1}' + z_{t-1} \dot{z}_{t-1}')(xG') + G \left(T^{-2} \sum_1^T z_{t-1} z_{t-1}' \right) \\ &\qquad + \left(T^{-2} \sum_1^T z_{t-1} z_{t-1}' \right) G' + T^{-1} \sum_1^T (\dot{z}_{t-1} u_t' + u_t \dot{z}_{t-1}') + O_p(T^{-1}). \end{aligned} \tag{21}$$

From Lemma 2.3 and (21) we now deduce that

$$\begin{aligned} &T^{-1} \sum_1^T (\dot{z}_{t-1} u_t' + u_t \dot{z}_{t-1}') \\ &\quad \Rightarrow G L_G(1, x)\, K_G(1, x)' + K_G(1, x)\, L_G(1, x)' G' \\ &\qquad - xG \left\{ G \int_0^1 L_G(r, x)\, K_G(r, x)'\, dr + \int_0^1 K_G(r, x)\, L_G(r, x)'\, dr G' \right\} \\ &\qquad - \left\{ G \int_0^1 L_G(r, x)\, K_G(r, x)'\, dx + \int_0^1 K_G(r, x)\, L_G(r, x)'\, dG' \right\} xG \\ &\qquad - G \int_0^1 K_G(r, x)\, K_G(r, x)'\, dr - \int_0^1 K_G(r, x)\, K_G(r, x)'\, dr G'. \end{aligned} \tag{22}$$

Now let $C = xG$ in (19) and differentiating (19) we have (noting that $J_{xG}(r) = K_G(r, x)$ and $(d/dx)J_{xG}(r) = GL_G(r, x)$)

$$\begin{aligned}
&GL_G(1, x)\,K_G(1, x)' + K_G(1, x)\,L_G(1, x)'G' \\
&\quad = G\int_0^1 K_G(r, x)\,LK_G(r, x)'\,dr + \int_0^1 K_G(r, x)\,K_G(r, x)'\,drG' \\
&\quad + xG\left\{G\int_0^1 L_G(r, x)\,K_G(r, x)'\,dr + \int_0^1 K_G(r, x)\,L_G(r, x)\,drG'\right\} \\
&\quad + \left\{G\int_0^1 L_G(r, x)\,K_G(r, x)'\,dr + \int_0^1 K_G(r, x)\,L_G(r, x)\,drG'\right\}(xG') \\
&\quad + G\int_0^1 L_G(r, x)\,dB(r)' + \int_0^1 [dB(r)\,L_G(r, x)']\,G'. \qquad (23)
\end{aligned}$$

It follows from (22) and (23) that

$$T^{-1}\sum_1^T (\dot{z}_{t-1}u_t' + u_t\dot{z}_{t-1}) \Rightarrow G\int_0^1 L_G(r, x)\,dB(r)' + \int_0^1 [dB(r)\,L_G(r, x)']\,G'$$

as required for part (a).

To prove part (b) we note first from (18) that $\dot{z}_t = Gw_t$, where

$$w_t = \sum_1^t \exp\{((t-j)\,x/T)G\}(t-j)/Tu_j + (t/T)\exp\{tx/T)G\}\,y_0.$$

Thus, from part (a) we have

$$\begin{aligned}
&G\left(T^{-1}\sum_1^T w_{t-1}u_t'\right) + \left(T^{-1}\sum_1^T u_t w_{t-1}'\right)G' \\
&\qquad \Rightarrow G\int_0^1 L_G(r, x)\,dB(r)' + \int_0^1 dB(r)\,L_G(r, x)'G'.
\end{aligned}$$

It follows that

$$\operatorname{tr}\left\{G\left(T^{-1}\sum_1^T w_{t-1}u_t'\right)\right\} \Rightarrow \operatorname{tr}\left\{G\int_0^1 L_G(r, x)\,dB(r)'\right\}. \qquad (24)$$

Since (24) holds for all matrices G we deduce that

$$T^{-1}\sum_1^T w_{t-1}u_t' \Rightarrow \int_0^1 L_G(r, x)\,dB(r)'.$$

Result (b) follows directly.

To prove (c) we integrate with respect to x over the interval $[0, l]$. We have

$$T^{-1}\sum_{1}^{T}\int_{0}^{l}\dot{z}_{t-1}(x)u_t'\,dx = T^{-1}\sum_{1}^{T} z_{t-1}(l)u_t' - T^{-1}\sum_{1}^{T} y_{t-1}u_t'$$

and

$$\int_0^l G\int_0^1 L_G(r,x)\,dB(r)'\,dx = \int_0^1\int_0^l GL_G(r,x)\,dx\,dB(r)'$$

$$= \int_0^1 K_G(r,l)\,dB(r)' - \int_0^1 B(r)\,dB(r)'.$$

Part (c) now follows from (b) and the continuous mapping theorem.

Proof of Theorem 2.6. We work from part (c) of Lemma 2.5. First let $G = fI_n$ for some $f<0$ and write

$$T^{-1}\sum_{1}^{T} z_{t-1}(l)u_t' = e^{-lf/T}T^{-1}\sum_{2}^{T}\left(\sum_{1}^{t-1} e^{(t-j)lf/T}u_j\right)u_t'$$

$$= e^{-lf/T}\sum_{s=1}^{T-1} e^{slf/T}\left(T^{-1}\sum_{t=s+1}^{T} u_{t-s}u_t'\right). \tag{25}$$

Now let $l = T/M$. We shall allow $M\uparrow\infty$ as $T\uparrow\infty$ in such a way that $M/T\downarrow 0$ (and, thus, $l\uparrow\infty$). Equation (25) becomes

$$e^{-f/M}\sum_{s=1}^{T-1} e^{sf/M}\left(T^{-1}\sum_{t=s+1}^{T} u_{t-s}u_t'\right).$$

But $e^{-f/M}\to 1$ as $M\uparrow\infty$ and

$$\sum_{s=1}^{T-1} e^{sf/M}\left(T^{-1}\sum_{t=s+1}^{T} u_{t-s}u_t'\right)\xrightarrow[p]{}\Omega_1. \tag{26}$$

In fact, (26) is simply the Abel estimate of the component Ω_1 of the scaled spectral density matrix $\Omega = 2\pi f_{uu}(0)$ at the origin (see, e.g., Hannan [5, p. 279]). We deduce that

$$T^{-1}\sum_{1}^{T} z_{t-1}(l)u_t' - T^{-1}\sum_{1}^{T} y_{t-1}u_t'$$

has the same asymptotic distribution as $T\uparrow\infty$ (with $l = T/M\uparrow\infty$) as

$$\Omega_1 - T^{-1}\sum^{T} y_{t-1}u_t'. \tag{27}$$

Now consider

$$K_G(r, l) = \int_0^r e^{(r-s)lf}\,dB(s) \equiv N\left(0, \int_0^r e^{2(r-s)lf}\,ds\,\Omega\right)$$
$$\equiv N(0, ((e^{2rlf} - 1)/2lf)\Omega).$$

Since $f < 0$ we deduce that

$$K_G(r, l) \xrightarrow[p]{} 0$$

as $l \uparrow \infty$. We may also show that

$$\int_0^1 K_G(r, l)\,dB(r)' \xrightarrow[p]{} 0. \tag{28}$$

Now part (c) of Lemma 2.5 holds for all l, so that combining (27) and (28) with part (c) we obtain

$$T^{-1} \sum_1^T y_{t-1} u_t' - \Omega_1 \Rightarrow \int_0^1 B(r)\,dB(r)',$$

giving the result as stated. Note also that

$$\begin{aligned}\lim_{T\to\infty} T^{-1} \sum_1^T y_{t-1} u_t' &= \lim_{T\to\infty} T^{-1} \sum_2^T \sum_{s=1}^{t-1} E(u_{t-s} u_t') \\ &= \lim_{T\to\infty} \sum_{s=1}^{T-1} (1 - s/T)\,E(u_1 u_{s+1}') \\ &= \sum_{k=2}^{\infty} E(u_1 u_k').\end{aligned}$$

Part (b) of the theorem follows in a similar way.

Proof of Theorem 2.8. This follows as a consequence of Theorem 2.6(a) and (6) since joint weak convergence of the numerator and denominator matrices in the quotient defining $T(\hat{A} - I)$ holds and an application of the continuous mapping theorem yields (20) as stated.

References

[1] Berkes, I., and Philipp, W. (1981). Approximation theorems for independent and weakly dependent random vectors. *Ann. Probab.* 7, 29–54.

[2] Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[3] Box, G. E. P., and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control.* Holden-Day, San Francisco.

[4] Eberlain, E. (1986). On strong invariance principles under dependence assumptions. *Ann. Probab.* **14**, 260–270.
[5] Hannan, E. J. (1970). *Multiple Time Series.* Wiley, New York.
[6] Ibragimov, I. A., and Linnik, Y. V. (1971). *Independent and Stationary Sequences of Random Variables.* Wolters–Noordhoff, Groningen.
[7] McLeish, D. L. (1975). A maximal inequality and dependent strong laws. *Ann. Probab.* **3**, 829–839.
[8] Philipp, W., and Stout, W. (1975). Almost sure invariance principles for weakly dependent random variables. *Mem. Amer. Math. Soc.* **161**, 1–140.
[9] Phillips, P. C. B. (1988). Regression theory for near-integrated time series. *Econometrica,* (in press).
[10] Phillips, P. C. B. (1986). Time series regression with a unit root. *Econometrica* **55**, 277–301.
[11] Phillips, P. C. B. (1986). Asymptotic expansions in nonstationary vector autoregressions. *Econometric Theory* **3**, 45–68.
[12] Phillips, P. C. B. (1987). Towards a unified asymptotic theory for autoregression. *Biometrika* **74**, 535–547.
[13] Phillips, P. C. B., and Durlauf, S. N. (1986). Multiple time series regression with integrated processes. *Rev. Econom. Stud.* **53**, 473–495.
[14] White, J. S. (1958). The limiting distribution of the serial correlation coefficient in the explosive case. *Ann. Math. Statist.* **29**, 1188–1197.